

Marked up version showing changes made**In the Specification:****Paragraph starting at page 7, line 22:**

Refer now to Figure 1, which is a block diagram of a hosting service 13 for hosting a plurality of domains on the Internet 10. To simplify the following discussion, the preferred embodiments of the present invention will be described in terms of a hosting service having a plurality of computers 14, referred to as "nodes" or "servers". Each node services one or more domain names and has an IP address associated with that node. Each node includes a processor 15 having a RAM disk cache 16 for facilitating access to files stored on a disk [14] 17. The mapping of IP addresses to domain names is assumed to be handled by DNS 12, which provides the user 10 with the IP address of the node in host 13 that is to service that user's request. If a site is too large to be serviced by a single node, the site is replicated on a plurality of nodes, and the DNS server uses an algorithm such as the round-robin algorithm discussed above to distribute requests for that site's files across the relevant nodes.

In the Claims:

10(Amended). The method of Claim [7] 1 wherein said router selects which of said service nodes corresponding to said two groups will service a request for that site.

REMARKS

The Examiner objected to the drawings because "files stored on a disk 14" (on page 7, line 27 of the specification) should be "files stored on a disk 17" to comply with Figure 1. The above amendment to the specification corrects this defect, and hence, no new drawings are required.

The Examiner rejected Claims 1-12 under 35 U.S.C. 103(a) as being unpatentable over Yu, (US Patent No. 6,351,775) and further in view of Desai, (US Patent No. 6,434,608). Applicant traverses this rejection. To sustain a rejection under 35 U.S.C. 103 based on a combination of references, the Examiner must show that each element of the rejected claim is taught in at least one of the cited references. "The mere fact that a reference could be modified to produce the patented invention would not make the modification obvious unless it is suggested by the prior art." (*Libbey-Owens-Ford v. BOC Group*, 4 USPQ 2d 1097, 1103). "When the PTO asserts that there is an explicit or implicit teaching or suggestion in the prior art, it must indicate where such a teaching or suggestion appears in the reference" (*In re Rijckaert*, 28 USPQ2d, 1955, 1957).

With respect to Claims 1, 8, 9, 10, 11, and the claims dependent therefrom, Claim 1 contains two limitations that are relevant to this rejection:

(1) measuring the computational resources required to service said requests to each of said sites over a first time period; and

(2) grouping said sites into N groups, each group being assigned to a corresponding one of said server nodes such that for each pair of groups, the difference in the sum of said measured computational resources is within a first predetermined error value;

In making this rejection, the Examiner looks to col. 7, lines 18-37 of Yu as teaching limitation (2). The cited passage does not provide any information on the basis used to group the sites. The cited passage merely discusses the PICS protocol used to transmit information about changes in the mapping of the sites to servers.

Yu teaches the grouping of sites into classes and the assignment of each class to a server (Abstract). In the scheme taught in Yu each server may be assigned a plurality of classes. Yu makes the assignment of the classes by collecting statistics on the number of times a request is made for an object in a given class from a given server, $CS(i,j)$, the total number of requests for all objects on a given server, $SA(j)$, and the total number of requests for objects in a each class $CA(j)$ (column 9, lines 27-33). Yu does not measure the resources or any other parameter relating to the individual elements of the classes; Yu measures only the number of requests for the various groups of sites, i.e., classes. Hence, Yu could not teach limitation (2).

The Examiner admits that Yu does not teach limitation (1). The Examiner looks to Desai for this missing teaching. The Examiner specifically points to column 3, lines 14-44 of Desai as teaching a system in which the number of cache misses that occur when a an object is retrieved are recorded. Desai then goes on to teach a system in which an object is only cached if previously received requests for that object generated in a given number of cache misses.

First, Applicant must point out that the Examiner's argument depends on there being a correlation between the number of cache misses and the computational resources required to service a request for the object generating the miss. The Examiner has not pointed to any teaching in the art that such a correlation exists. A cache miss occurs when the object being requested is not in the cache. The absence of the object from the cache can be the result of it not ever having been in the cache, or the result of the object having been replaced by an object that was requested subsequent to the last time the object was in the cache. Neither of these situations is correlated with the size or computational resources of the object. Hence, Applicant submits that there is no such correlation.

Second, the combination of the teachings of Yu and Desai is not a system that satisfies the limitations of Claim 1 or the claims dependent therefrom. The combination of these teachings would be a partitioning system as taught in Yu wherein the servers that provide the requested objects selectively cache objects in a cache based on the frequency of cache misses. In this regard, it should be noted that the frequency of cache misses is determined by the collection of objects on the server, not by the resources associated with a

given object. It is the interaction and timing of the requests that generates the cache misses. Hence, it is unclear how one would partition the classes among the servers based on cache misses as observed on a prior partitioning of the classes.

With respect to Claim 3, the Examiner points to Column 3, lines 14-44 and column 6, line 62 to column 7, line 15 as teaching the additional limitation introduced by this claim, namely, that

said sites return files in response to said requests, and wherein said step of measuring said computational resources comprises recording information identifying each returned file, the size of that file, and the number of times that file was returned.

Applicant disagrees with the Examiner's reading of these two passages. The first passage does not contain any reference to measuring file size for any purpose. The second passage refers to comparing the requested file size to the space in the cache of the server that is to supply the file. If there is sufficient space in that cache, the requested file is cached independent of whether or not it generated cache misses in the past. If there is insufficient space in that cache, the file is only cached if it has generated the required number of cache misses in the past. There is no teaching in the cited passage of recording information identifying the returned file or the number of times that file has been returned. Accordingly, there are additional grounds for allowing Claim 3 and the claims dependent therefrom.

With reference to Claim 4, the claim introduces the further limitation that

said step of grouping said sites also depends on the amount of memory in said cache memory on each of said servers.

The Examiner maintains that Yu teaches grouping the sites in a manner that also depends on the amount of memory in the cache memory on each of the servers. The Examiner cites col. 8, line 10 – col. 9, line 10 of Yu in support of this contention. Applicant must disagree with the Examiner's reading of Yu. The cited passage refers to the method by which the decision to place an object in a cache memory on a server is made. There is no teaching of using the size of the cache memories in determining the manner in which the classes are partitioned

between the servers. Accordingly, there are additional grounds for allowing Claim 4 and the claims dependent therefrom.

With reference to Claim 5, the claim introduces the additional limitation that

said groups are chosen such that said files returned during said first time period more than a predetermined number of times can be stored simultaneously in said cache memory.

The Examiner states that Desai teaches this limitation in the passage at col. 3, lines 14-44. The cited passage refers to how the decision to cache or not cache an object is made. The decision is based on the number of times a request for the object generated a cache miss in a previous time interval not on whether objects that were returned more than a predetermined number of times in a time interval will fit into the cache. In fact, as noted above, Desai does not teach recording the number of times an object is returned. Desai only teaches recording of the number of times a request for an object generates a cache miss. Accordingly, there are additional grounds for allowing Claim 5 and the claims dependent therefrom.

With reference to Claim 6, this claim introduces the additional limitation that

said measurement of said computational resources further comprises measuring the number of bytes of data returned in response to said requests for each site during said first time period.

The Examiner maintains that the passage from col. 6, lines 62 to col. 7 line 1 of Desai teaches this limitation. As noted above, this passage teaches referring to the size of an object in determining whether or not to cache that object. It does not teach measuring the number of bytes of data returned in services requests for that object. Accordingly, there are additional grounds for allowing Claim 6 and the claims dependent therefrom.

With reference to Claim 7, this claim introduces the additional limitation that the step of measuring the resources further includes

estimating the number of bytes of data returned directly from said cache memory in servicing said requests for each site during said first time period.

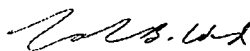
The Examiner maintains that the passage from col. 6, lines 62 to col. 7 line 10 of Desai teaches this limitation. As noted above, this passage teaches referring to the size of an object in determining whether or not to cache that object. It does not teach measuring the number of bytes of data returned directly from the cache in servicing requests for that object. Accordingly, there are additional grounds for allowing Claim 7 and the claims dependent therefrom.

With reference to Claim 12, this claim introduces the limitation that groups are moved from one server to the other in a repartitioning operation in a manner that minimizes the number of swaps that must be done. The Examiner maintains that Yu teaches this in the passage at col. 10, lines 25-42. Applicant must disagree with the Examiner's reading of Yu. The passage in question teaches that a class may be moved from one server to another or to multiple servers. However, there is no teaching that the repartitioning depends on the number of changes that are to be made in the original partition. Accordingly, there are additional grounds for allowing Claim 12.

The above amendment to the claims corrects a typographical error.

I hereby certify that this paper is being sent by FAX to 703-746-7239.

Respectfully Submitted,



Calvin B. Ward
Registration No. 30,896
Date: February 3, 2003

IP Administration
Legal Dept. MS 35
HEWLETT PACKARD CO.
P.O. Box 272400
Fort Collins, CO 80528-9599